

## **DWR Data Standards Subcommittee**

### **Comments on Draft State Minimum Metadata Standards**

**Version: 2012-05-09**

If there is no specific comment below, we agreed with the recommended standard item.

#### **Date:**

The date or range of dates when the data were gathered, or the date the photos, maps or other items at the core of the data set, were created.

Our discussion noted that referring to date by either when data were gathered or when they were created could lead to confusion about which it is. Furthermore, “date” could also mean when the data was published, modified, or copied. The FGDC in Section 1.3 has two pieces of information here, not one. The date/time information, and the “currentness reference”. Perhaps the date should be *\*recommended\** as publication date but not required.

Also, “Year”, rather than “date”, may be more appropriate for legacy data. Figuring out the date(s) data were gathered may be impossible for legacy data.

#### **Field Definitions (List & Define each field):**

DWR had some discussion about this. Some felt it was onerous, others felt it is absolutely necessary. If this is included, units should be defined for fields containing numeric values.

#### **Access Constraints:**

There should be a standard domain list of access constraint types.

**Use Constraints:**

This item generated some debate. Some feel this is useless, as it is unenforceable. Others feel:

“It’s not a matter of preventing people from misusing the data; this is a way to identify what system requirements need to be applied towards access. If something is public domain, then it can be served out via a public facing portal or web service. If it is for DWR internal use, then it will only be accessible via our department’s intranet. If it is restricted to a select audience, then special login permissions need to be in place. Metadata doesn’t accomplish data access control, it simply describes what level of control should be applied.”

**Distribution:**

Distribution is section 6 in the FGDC standards. What is asked for here appears to be only Section 6.4.2.2 - 6.4.2.2.1.1.1.1. We need the format of the information, Section 6.4.2.1.1, and version number, Section 6.4.2.1.2. This is the one piece of information that the proposal is missing.

**Progress:**

“Complete” and “incomplete” is not a sufficient list. As with access constraints, a full domain list should be developed and adhered to. In general, domain values for metadata elements should be used and provided whenever possible, not just the field requirements.

**ISO Topic Category:**

DWR staff had some discussion on this. It was pointed out that more than one category could apply. FGDC recommends using “at least one” topic category. However, some felt that even if that’s the case, that only one numeric ISO topic should be applied as category, because the whole point of categorizing is to formalize a dataset into one thematic category. DWR had not developed consensus on this subject at writing.

**Keywords:**

As with ISO Topic category, more than one ISO topic category number could apply. Some felt that, again, putting numerous ISO Topic #s creates more confusion than is worth. The alphabetic keywords themselves could include multiple ITO Topic names but should not include multiple Topic #s.

Again, a standard list of keywords should be used if possible.

Another concern is that a data steward may just put in one or two keywords to satisfy this requirement, and the keywords would be repetitive of the ISO themes. The stewards should therefore be instructed/encouraged to branch out beyond just ISO themes. State Program name is also a good idea, and year of publishing would also be useful. A standard place keyword set should be developed and used.

Others felt that these were good suggestions, but might not qualify as requirements, since these could include legacy and external data.

The FGDC recommendation is to use “controlled vocabularies (thesauri).